Accepted for publication in *Journal of Engineering Mechanics, Trans. ASCE*

May 2014

Bayesian Updating with Structural Reliability Methods

Daniel Straub, Iason Papaioannou

Engineering Risk Analysis Group, Technische Universität München (straub@tum.de)

Abstract

Bayesian updating is a powerful method to learn and calibrate models with data and observations. Because of the difficulties involved in computing the high-dimensional integrals necessary for Bayesian updating, Markov Chain Monte Carlo (MCMC) sampling methods have been developed and successfully applied for this task. The disadvantage of MCMC methods is the difficulty of ensuring the stationarity of the Markov chain. We present an alternative to MCMC that is particularly effective for updating mechanical and other computational models, termed BUS: Bayesian Updating with Structural reliability methods. With BUS, structural reliability methods are applied to compute the posterior distribution of uncertain model parameters and model outputs in general. We propose an algorithm for the implementation of BUS, which can be interpreted as an enhancement of the classical rejection sampling algorithm for Bayesian updating. This algorithm is based on the subset simulation and its efficiency is not dependent on the number of random variables in the model. The method is demonstrated by application to parameter identification in a dynamic system, Bayesian updating of the material parameters of a structural system, and Bayesian updating of a random-field-based FE model of a geotechnical site.

Keywords

Bayesian updating; structural reliability; sampling; measurements; monitoring; FEM.

Introduction

With advances in information and sensor technology, increasing amounts of data on engineering systems are collected and stored; examples include data on deformations and dynamic properties of structural systems, and data on ambient factors influencing deterioration of engineering structures. This information can – and should – be used to reduce the uncertainty in engineering models and optimize the management of these systems. As an example, a smart structure should use sensor information to automatically trigger actions like detailed inspections or system shut-downs.

A consistent and effective framework for combining new information with existing models is provided by Bayesian analysis, in which prior probabilistic models are updated with data and observations. The Bayesian framework enables the combination of uncertain and incomplete information with models from different sources and it provides probabilistic information on the accuracy of the updated model. The latter is of particular relevance, since system predictions typically remain uncertain even with new information. For this reason, important decisions on engineering systems should be made on the basis of reliability and risk assessments, and Bayesian analysis is a cornerstone of such assessments.

Bayesian updating of engineering and mechanical models has been considered since the 1960s. Benjamin and Cornell (1970) described the use of Bayesian updating for improved engineering decision making through examples from material testing and geotechnical site investigation. Tang (1973) recognized the potential of Bayesian updating for updating the probabilistic description of material imperfections and flaws with inspection information, an idea that later formed the basis for reliability- and risk-based planning of inspections using Bayesian principles (Yang and Trapp 1974; Straub and Faber 2005). Bayesian analysis has also been used extensively for structural identification, i.e. the task of identifying dynamic properties of structural systems based on vibration measurements (Natke 1988; Beck and Katafygiotis 1998). In hydrology, Bayesian analysis has been frequently applied for model calibration with measurements, e.g. of rainfall and discharge measurements (Kavetski et al. 2006; Beven 2008). The topic has also attracted the attention of the mathematical community (Kennedy and O'Hagan 2001). Overall, countless applications of Bayesian updating of mechanical, engineering and computational models are reported in the literature. Its popularity is further increasing as computational limitations are becoming less of a concern due to increased computing power and enhanced algorithms.

Bayesian updating requires the evaluation of the posterior probabilistic model given the prior model and the likelihood function describing observation data. With few exceptions, the posterior model must be evaluated numerically. For cases where the information contained in the data is much stronger than the prior model, the posterior can be approximated in terms of an asymptotic expression (Beck and Katafygiotis 1998; Papadimitriou et al. 2001). However, this expression requires determination of the possibly multiple local maxima of the likelihood function as well as evaluation of the Hessian of the likelihood at each corresponding parameter set. Most commonly, Bayesian updating is performed through sampling methods. The Markov Chain Monte Carlo (MCMC) method is particularly popular, since it allows direct sampling from the posterior distribution without the need to solve the potentially high-dimensional integral in the Bayesian formulation (Gilks et al. 1998; Gelman 2004). Many authors have applied and adopted MCMC to Bayesian updating of mechanical models, including (Beck and Au 2002; Cheung and Beck 2009; Sundar and Manohar 2013). A main problem of MCMC methods is that it cannot generally be ensured that the samples have reached the stationary distribution of the Markov chain, i.e. the posterior distribution (Plummer et al. 2006). Various alternatives to MCMC exist, which are mostly based on rejection sampling and importance sampling, e.g. the adaptive rejection sampling from a log-concave envelope distribution, which is effective for updating single random variables (Gilks and Wild 1992), or generalized sequential particle filter methods for updating arbitrary static or dynamic systems (Chopin 2002; Ching and Chen 2007). These methods are often combined with MCMC.

An increasingly popular approach to represent probabilistic models in engineering is the Bayesian network (BN) framework, which allows decomposing joint probability distributions into local conditional distributions. As its name suggests, the BN modeling framework is particularly suitable for Bayesian updating. Under certain conditions, mechanical models can be effectively represented in a BN with discrete or discretized random variables and efficient algorithms are then available (Straub and Der Kiureghian 2010b). Examples of BN applied to updating of mechanical models are presented e.g. in (Mahadevan et al. 2001; Straub 2009; Straub and Der Kiureghian 2010a). The BN framework may also be combined with continuous random variables, e.g. it is ideally suited for applications of MCMC algorithms based on Gibbs sampling (Gilks et al. 1994). Unfortunately, the conditions for the effectiveness of BN, namely conditional independence among random variables, are often not given in mechanical models.

In this paper, a novel approach to Bayesian updating of engineering models is presented, which combines a rejection sampling strategy with structural reliability methods. We term the method BUS (Bayesian Updating with Structural reliability methods). BUS is based on the approach

described in (Straub 2011) for Bayesian updating of the reliability of engineering systems. In this paper, the approach is extended to computing the posterior distribution of uncertain model parameters and model outputs in general. Key advantages of the BUS approach are its simplicity and that it enables using the whole portfolio of structural reliability methods and associated software for Bayesian updating. Here, we implement the BUS approach through subset simulation (Au and Beck 2001), which can efficiently handle models with many random variables. Three illustrative examples are included, which demonstrate the application of the method to structural identification, Bayesian updating of material parameters in a statically loaded structure, and updating of a random-field-based finite-element (FE) model.

Bayesian updating of mechanical models

Mechanical models consist of a set of equations and boundary conditions that describe the geometry, material properties and loading conditions. In many applications, some of the parameters of the model are uncertain. These parameters are modeled as random variables **X**, characterized through their joint probability density function (PDF) $f(\mathbf{x})$. The mechanical model itself is often uncertain, which can be reflected by introducing one or more additional random variables into **X** to represent model errors.

When new observations or data related to a mechanical model are available, these can be used to learn the mechanical model. As an example, observations of deformations of a structure under known loading provide information on the material properties of the structure. Since the uncertainty in the mechanical model and its parameters is reflected through **X**, learning from the observations is tantamount to updating the distribution of **X**. Bayes' rule enables updating a *prior* probability distribution $f'(\mathbf{x})$ with observations or data to a *posterior* probability distribution $f''(\mathbf{x})$:

$$f''(\mathbf{x}) = \frac{L(\mathbf{x})f'(\mathbf{x})}{\int_{\mathbf{X}} L(\mathbf{x})f'(\mathbf{x})\mathrm{d}\mathbf{x}}. \tag{1}$$

Throughout this paper we use the convention $\int_{\mathbf{X}} \mathrm{d}\mathbf{x} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \mathrm{d}x_1 \dots \mathrm{d}x_n$. The direct evaluation of the n-fold integral in Eq. (1) is not feasible in the general case, which has motivated the introduction of MCMC and other sampling techniques for Bayesian analysis.

$L(\mathbf{x})$ in Eq. (1) is the so-called likelihood function and describes observations and data (Fisher 1922). It is defined as

$$L(\mathbf{x}) \propto \Pr(Observation|\mathbf{X} = \mathbf{x}). \quad (2)$$

When observations are made, they often correspond to outcomes of the mechanical models. Therefore, the likelihood function must include the mechanical models to relate the observation to the model parameters $\mathbf{X}$. As an example, if deformations of a structure are measured, the model predictions of these deformations for given values of $\mathbf{X}$ are required. Let $h_i(\mathbf{X})$ denote such a model prediction. Furthermore, let y_i denote the corresponding observed deformation and let ϵ_i denote the deviation of the model prediction from the observation. This deviation is due to measurement errors and model errors (those not modeled explicitly through $\mathbf{X}$); it is modelled through the PDF $f_{\epsilon_i}(\cdot)$. The following relationship holds: $y_i - h_i(\mathbf{X}) = \epsilon_i$. The likelihood function $L_i(\mathbf{x})$ describing this observation is therefore

$$L_i(\mathbf{x}) = f_{\epsilon_i}[y_i - h_i(\mathbf{x})]. \quad (3)$$

Generally, one can distinguish two classes of observations (Madsen et al. 1985; Straub 2011):

(a) Observations providing equality information. These are measurements y_i of continuous quantities, for which equalities similar to $y_i - h_i(\mathbf{X}) = \epsilon_i$ can be formulated. The likelihood function for these observations is

$$L_i(\mathbf{x}) = f_{Y_i|\mathbf{X}}(y_i|\mathbf{x}) \ . \quad (4)$$

$f_{Y_i|\mathbf{X}}$ is the PDF of the measured quantity given $\mathbf{X} = \mathbf{x}$. Typically, $f_{Y_i|\mathbf{X}}$ is defined by the PDF of the observation error ϵ_i, which is often assumed to follow a Gaussian distribution. In case of an additive observation error ϵ_i, $L_i(\mathbf{x})$ is defined as in Eq. (3). In case of a multiplicative observation error ϵ_i, the relationship between measurement y_i and model prediction $h_i(\mathbf{X})$ is $y_i = h_i(\mathbf{X}) \cdot \epsilon_i$; solving for ϵ_i and inserting into $f_{\epsilon_i}(\epsilon_i)$, the likelihood function is obtained as

$$L_i(\mathbf{x}) = f_{\epsilon_i}\left(\frac{y_i}{h_i(\mathbf{x})}\right). \quad (5)$$

(b) Observations providing inequality information. These are observations that are characterized by a finite probability of occurrence, e.g. observations of categorical values or observations of system performances such as failure/survival, but also censored data. Commonly, it is possible to formulate a model outcome $h_i(\mathbf{X})$ such that the observation event is defined through

$$Z_i = \{\mathbf{x} \in \mathbb{R}^n : h_i(\mathbf{x}) \leq 0\}. \quad (6)$$

A model outcome $h_i(\mathbf{X})$ of this form is known in structural reliability as *limit state function*. The corresponding likelihood function is then

$$L_i(\mathbf{x}) = \Pr(Z_i | \mathbf{X} = \mathbf{x}) = I[h_i(\mathbf{x}) \leq 0]. \quad (7)$$

I is the indicator function. This likelihood thus takes on values 0 or 1.

In the general case, m observations are available, of which m_y are of the equality type and m_z are of the inequality type. Under the common assumption that the measurement/observations are statistically independent given the model parameters $\mathbf{x}$, the combined likelihood of all observations is

$$L(\mathbf{x}) = \prod_{i=1}^{m} L_i(\mathbf{x}). \quad (8)$$

In case of statistically dependent observation errors ϵ_i, the combined likelihood must be formulated as a function of the joint PDF of all ϵ_i. An example of statistically dependent observation errors is included in the applications presented later in the paper.

Through Eq. (8), all observations are combined in the single expression $L(\mathbf{x})$. For the computations of the posterior distribution presented in the remainder of the paper, it is irrelevant whether the observations are of the equality type, the inequality type, or combinations thereof.

It should be clear from the preceding discussion that every computation of the likelihood function requires up to m model evaluations to determine $h_i,\ i = 1, \dots, m$. Luckily, for most applications, one model call is sufficient to compute all h_i, so the necessary number of model evaluations is equal to the number of likelihood function calls. For advanced numerical models, such as the geotechnical finite element model presented later in the numerical applications, these evaluations are computationally costly. Our aim is therefore to perform Bayesian updating, Eq. (1), with a minimum number of likelihood function calls.

Sampling algorithm

The goal is an efficient algorithm for sampling from the posterior distribution $f''(\mathbf{x})$. If direct evaluation of the integral in Eq. (1) is not feasible, the posterior PDF is known only up to a proportionality constant, i.e.

$$f''(\mathbf{x}) \propto L(\mathbf{x}) f'(\mathbf{x}). \tag{9}$$

Samples from $f''(\mathbf{x})$ can be generated through a simple rejection sampling algorithm, which is introduced in the following.

Let P be a standard uniform random variable in [0,1]. Consider the augmented outcome space $[\mathbf{x}; p]$ and define the domain

$$\Omega = \{p \leq cL(\mathbf{x})\}. \tag{10}$$

where c is a positive constant that ensures $cL(\mathbf{x}) \leq 1$ for all $\mathbf{x}$. The selection of this constant c is discussed in Annex A.

The posterior distribution of Eq. (1) can be written as

$$f''(\mathbf{x}) = \frac{\int_{p\in\Omega} f'(\mathbf{x})\,\mathrm{d}p}{\int_{[\mathbf{x},p]\in\Omega} f'(\mathbf{x})\,\mathrm{d}p\,\mathrm{d}\mathbf{x}} = \frac{\int_0^1 I([\mathbf{x},p]\in\Omega) f'(\mathbf{x})\,\mathrm{d}p}{\int_{\mathbf{X}}\int_0^1 I([\mathbf{x},p]\in\Omega) f'(\mathbf{x})\mathrm{d}p\,\mathrm{d}\mathbf{x}}. \tag{11}$$

The validity of this result can be demonstrated as follows. The denominator in (11) corresponds to a structural reliability problem with limit state function $h(\mathbf{x}, p) = p - cL(\mathbf{x})$,. It is equal to the probability of being in Ω under the prior PDF $f'(\mathbf{x})$, i.e.

$$\begin{aligned}\int_{\mathbf{X}}\int_0^1 I([\mathbf{x},p]\in\Omega) f'(\mathbf{x})\mathrm{d}p\,\mathrm{d}\mathbf{x} &= \int_{\mathbf{X}}\left\{\int_0^1 I[p\leq cL(\mathbf{x})]\mathrm{d}p\right\} f'(\mathbf{x})\mathrm{d}\mathbf{x} \\ &= \int_{\mathbf{X}} cL(\mathbf{x}) f'(\mathbf{x})\,\mathrm{d}\mathbf{x}\end{aligned} \tag{12}$$

The numerator in (11) is equal to

$$\int_{p\in\Omega} f'(\mathbf{x})\,\mathrm{d}p = \int_0^{cL(\mathbf{x})} f'(\mathbf{x})\,\mathrm{d}p = cL(\mathbf{x}) f'(\mathbf{x}) \tag{13}$$

Inserting (12) and (13) in (11), the original Bayesian updating formulation of Eq. (1) is obtained.

The posterior cumulative distribution function (CDF) is obtained by integrating Equation (11) on both sides:

$$F''(\mathbf{x}_0) = \int_{-\infty}^{\mathbf{x}_0} f''(\mathbf{x})\mathrm{d}\mathbf{x} = \frac{\int_{-\infty}^{\mathbf{x}_0}\left(\int_{p\in\Omega} f'(\mathbf{x})\,\mathrm{d}p\right)\mathrm{d}\mathbf{x}}{\int_{[\mathbf{x},p]\in\Omega} f'(\mathbf{x})\,\mathrm{d}p\,\mathrm{d}\mathbf{x}} \tag{14}$$

Generating samples $\mathbf{x}^{(k)}$ from $f'(\mathbf{x})$ and samples $p^{(k)}$ from the standard uniform distribution, $k = 1, \dots, K$, the following Monte Carlo approximation to the posterior CDF is obtained:

$$F''(\mathbf{x}_0) \approx \frac{\sum_{k=1}^{K} I(\mathbf{x}^{(k)} \le \mathbf{x}_0) I\left([\mathbf{x}^{(k)}, p^{(k)}] \in \Omega\right)}{\sum_{k=1}^{K} I([\mathbf{x}^{(k)}, p^{(k)}] \in \Omega)} \tag{15}$$

It follows that samples $\mathbf{x}^{(k)}$ generated from $f'(\mathbf{x})$ and falling into the domain Ω are distributed according to the posterior $f''(\mathbf{x})$. This leads to the following rejection sampling algorithm, which generates K samples of the posterior $f''(\mathbf{x})$. Note that this algorithm is equivalent to a classical rejection sampling where the prior distribution is applied as an envelope distribution and the likelihood as a filter (Smith and Gelfand 1992).

Simple rejection sampling algorithm

1. $k = 1$.
2. Generate a sample $\mathbf{x}^{(k)}$ from $f'(\mathbf{x})$.
3. Generate a sample $p^{(k)}$ from the standard uniform distribution in [0,1].
4. If $[\mathbf{x}^{(k)}, p^{(k)}] \in \Omega$
    a. Accept $\mathbf{x}^{(k)}$
    b. $k = k + 1$
5. Stop if $k = K$, else go to 2.

The rate of acceptance is equal to the denominator in Eq. (11),

$$p_{acc} = \int_{[\mathbf{x},p]\in\Omega} f'(\mathbf{x})\,\mathrm{d}p\,\mathrm{d}\mathbf{x} = \int_{\mathbf{X}} cL(\mathbf{x})f'(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{16}$$

For a fixed number of initial samples N_I, the number of accepted samples is binomial distributed with parameters N_I and p_{acc}. To generate a fixed number of accepted samples K, as in the above algorithm, the number of samples that must be generated N has the negative binomial distribution:

$$p_N(n) = \binom{n-1}{K-1} p_{acc}^{K} (1 - p_{acc})^{(n-K)} \tag{17}$$

The mean value of N is $\mathrm{E}[N] = K/p_{acc}$. Consequently, the expected number of samples to generate, and thus the expected number of evaluations of $L(\mathbf{x})$, is proportional to $1/p_{acc}$.

Let us consider a simple example to clarify the simple rejection sampling algorithm and its limitations.

Illustration: Updating of a single random variable

Let X be a random variable with standard normal prior, i.e. $f'(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$. A measurement of X is made, resulting in a value of 2. The measurement is associated with an additive error ϵ, which is normal distributed with zero mean and standard deviation $\sigma_\epsilon = 0.5$. The likelihood function is thus $L(x) = \frac{1}{0.5\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-2)^2}{0.5^2}\right)$. The parameter c is chosen as $c = [\max L(x)]^{-1} = 0.5\sqrt{2\pi}$. For this example, an analytical solution can be obtained for comparison: The posterior is the normal distribution with mean 1.6 and standard deviation $\sqrt{0.2} = 0.45$.

The domain Ω, together with samples of X and P, is shown in Figure 1a. In Figure 1b, the empirical frequency plot of the accepted samples is shown, together with the exact posterior CDF $F''(x)$. The sample mean and standard deviation of the accepted samples are 1.62 and 0.45; these values are close to the true solution.

In this example, 25 out of 200 samples are accepted. The exact mean acceptance rate can be computed as $p_{acc} = \int_{-\infty}^{\infty} cL(x)f'(x)\mathrm{d}x = 0.09$. Although not ideal, a mean acceptance rate of $p_{acc} = 0.09$ would be acceptable for most problems. However, the acceptance rate decreases quickly with increasing number of observations m. If m measurements are made of X in the above example, all with independent identically distributed (iid) observation errors ϵ_i, the average acceptance rate is proportional to $\frac{1}{\sqrt{m}}$ (see Annex B).

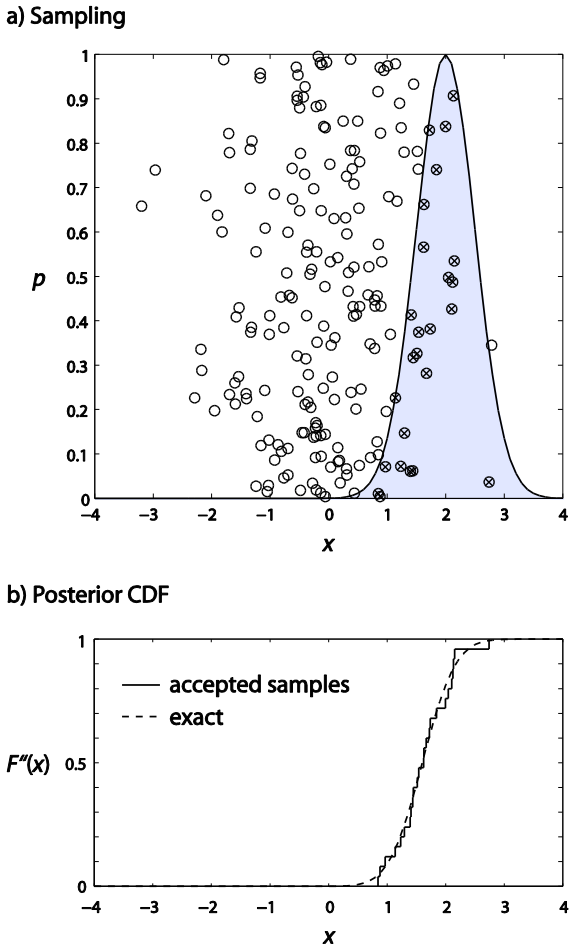

Figure 1. Illustration of the simple rejection sampling algorithm: Updating a single random variable X with the observation $X + \epsilon = 2$, using 200 samples.(a) Joint samples of X and P drawn from the prior distribution; accepted samples are indicated by a cross. The shaded area is the observation/acceptance domain Ω. (b) Resulting empirical frequency plot of the accepted samples, together with the analytical solution.

Note that the acceptance rate is directly proportional to the constant c of Eq. (10). Therefore, c should be selected as large as possible, while still ensuring that $cL(\mathbf{x}) \leq 1$ for all $\mathbf{x}$. In the above example, it is chosen as $c = [\max L(\mathbf{x})]^{-1}$, which is the optimal choice. However, in some real applications $\max L(\mathbf{x})$ can be difficult to evaluate. Strategies for selecting an optimal value of c in such cases are discussed in Annex A.

Bayesian updating with structural reliability methods (BUS)

The above rejection sampling algorithm quickly becomes inefficient with increasing number of observations due to the large rejection rate, unless the posterior $f''(\mathbf{x})$ is close to the prior $f'(\mathbf{x})$. This problem is inherent to any direct rejection sampling algorithm (Gelman 2004; Bolstad 2011) and has motivated the development of MCMC techniques for simulating from the posterior (Gilks et al. 1998). However, the advantage of the simple rejection sampling

algorithm over MCMC is that it is straightforward to implement and it is guaranteed to give exact, uncorrelated samples of the posterior. Therefore, in the following a method is proposed that maintains partly the advantages of the simple rejection sampling algorithm but has much higher efficiency. This method, which we call BUS, deals with the inefficiency problem of the rejection sampling by combining it with structural reliability methods.

As pointed out earlier, the integral in the denominator of Eq. (11) corresponds to a structural reliability problem. In structural reliability, the domain Ω describes the event of interest, typically a failure event; in the context of the Bayesian updating of Eq. (11), it describes an observation event, which we denote by Z_e. The integral, and therefore $\Pr(Z_e)$, is equal to the acceptance rate p_{acc} of the simple rejection sampling algorithm, which decreases with the square root of the number of observations. In many applications, $\Pr(Z_e)$ will be too small for Monte Carlo simulation to be efficient. What is needed are methods that allow to more efficiently explore the domain Ω and compute $\Pr(Z_e)$ even when it is very small. Structural reliability methods such as FORM/SORM or importance sampling methods have been specifically developed to evaluate small probabilities (Ditlevsen and Madsen 1996; Rackwitz 2001). They approximate either the limit state function or the PDF of $\mathbf{X}$ in the neighborhood of the domain Ω with the highest probability density, which in the context of Bayesian updating corresponds to the region of the highest posterior probability density. These methods are therefore ideally suited to enhance the efficiency of the rejection sampling algorithm.

When applying structural reliability methods, it is convenient to transform the problem from the outcome space of the original random variables P and $\mathbf{X}$ to a space with independent standard normal random variables $\mathbf{U} = [U_0; U_1; \dots; U_n] \in \mathbb{R}^{n+1}$. Since P and $\mathbf{X}$ are independent, they can be transformed separately. The transformation from $\mathbf{U}$ to P and $\mathbf{X}$ is

$$P = \Phi(U_0), \tag{18}$$

where $\Phi =$ standard normal CDF, and

$$\mathbf{X} = \mathbf{T}(U_1, \dots, U_n). \tag{19}$$

$\mathbf{T} =$ one of the classical transformations used in structural reliability methods. Both the Rosenblatt transformation (Hohenbichler and Rackwitz 1981) or the marginal transformation based on the Nataf model (Der Kiureghian and Liu 1986) can be applied.

The domain Ω describing the observation can now be transformed to an equivalent domain Ω_U in standard normal space:

$$\Omega_U = \{\Phi(u_0) \leq cL\big(\mathbf{T}(u_1, \dots, u_n)\big)\}. \tag{20}$$

For convenience, we define Ω_U in terms of a function H,

$$H(\mathbf{u}) = u_0 - \Phi^{-1}\Big(cL\big(\mathbf{T}(u_1, \dots, u_n)\big)\Big), \tag{21}$$

so that $\Omega_U = \{H(\mathbf{u}) \leq 0\}$. H is a limit state function describing the observations in standard normal space, as may be evident from the correspondence of $\{H(\mathbf{u}) \leq 0\}$ to the definition of the inequality observation in Eq. (6).

For illustration, Figure 2 shows the transformation of the observation domain Ω and the samples of $\mathbf{X}$ shown in Figure 1a to the corresponding domain Ω_U and samples in standard normal space. In order to obtain samples from the posterior, the samples falling into the domain Ω_U must be transformed to the original space using Eq. (19).

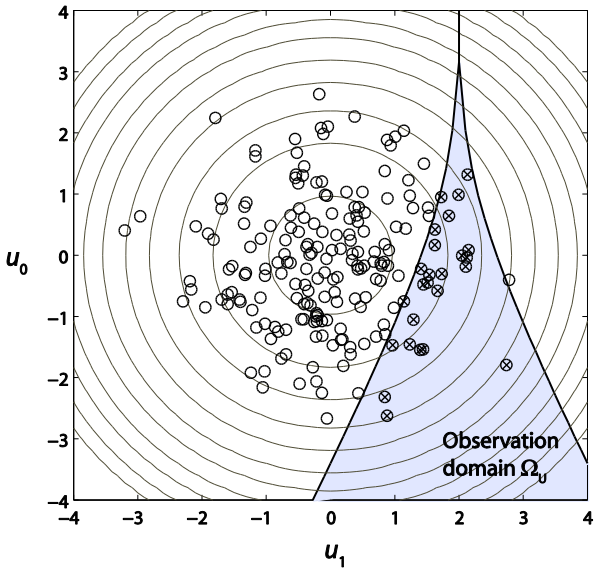

Figure 2. Observation domain and samples of Figure 1a transformed to standard normal space.

A main advantage of BUS is that it can potentially be combined with a large number of available structural reliability methods. It is not the intention of this paper to identify the most optimal method, since this choice will depend on problem as well as on the preferences and experience of the analyst. The strength of expressing the observations and data through limit state functions H is exactly the flexibility that it provides.

Among the available structural reliability methods, importance sampling (IS) methods are suitable for generating (weighted) samples from the observation domain Ω_U. Such IS methods

may be based on the design point (the Most Likely Failure Point MLFP) identified by means of the first-order reliability method (FORM). Note that in the context of Bayesian updating, the design point corresponds to the mode of the posterior distribution. Axis parallel IS based on the design point (also known as line sampling) was applied in an example presented in (Straub 2011) for the updating of the failure probability; it would also be efficient for updating the probability distribution of **X** using BUS. Since the identification of the design point is an optimization problem, alternative IS methods that do not require the design point may be preferable. In addition to classical adaptive IS methods (e.g. Bucher 1988), newly proposed approaches such as the one of (Kurtz and Song 2013) appear to be promising for the application with BUS. When applying IS concepts, the resulting samples are weighted. To obtain unweighted samples, a resampling scheme must be applied: In an additional step, independent samples are drawn from a discrete probability distribution defined through the original samples and their weights (Doucet et al. 2001).

In principle it is possible to use FORM (or SORM) in the context of the BUS approach. If FORM was applied, the posterior distribution in **U**-space would be approximated by a censored standard multi-normal distribution; the censoring is defined through the hyperplane given by $\nabla H(\mathbf{u}^*)^{\mathrm{T}}(\mathbf{u}-\mathbf{u}^*)=0$, where $\mathbf{u}^*$ is the design point and ∇H is the gradient vector of the observation limit state H of Eq. (21). However, the use of FORM or SORM is not advocated without further investigations on its accuracy. As can be observed from Figure 2, the shape of the observation domain differs from the failure domains typically encountered in structural reliability.

An alternative to FORM/SORM and importance sampling methods is subset simulation (SuS) proposed by (Au and Beck 2001), which is particularly efficient for structural reliability problems where the number of random variables n is large. For the considered implementation in BUS, the SuS has the additional advantage that in its final step it directly produces samples from the posterior distribution.

BUS algorithm based on subset simulation

In the following, we introduce an algorithm that uses SuS to generate samples in the observation domain Ω_U. It assumes that the observation event is transformed to **U**-space by means of Eq. (21). The algorithm is applied in the numerical examples. We restrict the presentation of SuS to a summary of its main principles. For details on its implementation in the examples presented later in the paper, the reader is referred to (Papaioannou et al. 2014).

SuS, originally proposed by Au and Beck (2001), evaluates the probability $\Pr(Z_e)$ of an event Z_e associated with a limit state function $H(\mathbf{u})$ and here defined as $Z_e = \{\mathbf{u} \in \mathbb{R}^{n+1}: H(\mathbf{u}) \leq 0\}$. SuS is based on expressing the event Z_e as the intersection of M intermediate events that are nested, i.e. it holds $Z_1 \supset Z_2 \supset \cdots \supset Z_M = Z_e$. The probability $\Pr(Z_e)$ is then expressed as

$$\Pr(Z_e) = \Pr\left(\bigcap_{i=1}^{M} Z_i\right) = \prod_{i=1}^{M} \Pr(Z_i|Z_{i-1}) \tag{22}$$

where Z_0 is the certain event. Thus the possibly small probability $\Pr(Z_e)$ is expressed as the product of larger conditional probabilities. The intermediate events Z_i, $i > 0$, are defined as $Z_i = \{\mathbf{u} \in \mathbb{R}^{n+1}: H(\mathbf{u}) \leq b_i\}$, where $b_1 > b_2 > \cdots > b_M = 0$. The values of b_i can be chosen adaptively, so that the estimates of the conditional probabilities correspond to a chosen value p_0. To this end, J samples are simulated from the random vector $\mathbf{U}$ conditional on each intermediate domain Z_{i-1}. For each sample, the limit state function $H(\mathbf{u})$ is evaluated and the samples are ordered in increasing order of magnitude of their limit state function values. The threshold b_i is then set to the p_0-percentile of the ordered samples. This procedure is repeated until the maximum level M is reached, for which $b_M = 0$. The samples conditional on the certain event Z_0 are obtained by crude Monte Carlo sampling. The samples conditional on the events Z_i, for $i = 1, \ldots, M-1$, are computed by simulating states of Markov chains through MCMC starting from the samples conditional on Z_{i-1} for which $H(\mathbf{u}) \leq b_i$. It is noted that the seeds of the Markov chains always follow the target distribution and hence the applied MCMC does not suffer a convergence (burn-in) problem (Zuev et al. 2012; Papaioannou et al. 2014). The probability of $\Pr(Z_e)$ can be approximated by:

$$\Pr(Z_e) \approx p_0^{M-1} \hat{P}_M, \tag{23}$$

where $\hat{P}_M$ is the estimate of the conditional probability $\Pr(Z_M|Z_{M-1})$ and is given by the ratio of the number of samples for which $H(\mathbf{u}) \leq 0$ over the total number of samples J simulated conditional on Z_{M-1}. The value p_0 of the intermediate probabilities and the number of samples J in each intermediate step are chosen by the analyst. Au and Beck (2001) suggested using $p_0 = 0.1$, whereas Zuev et al. (2012) showed that a choice of $p_0 \in [0.1, 0.3]$ leads to similar efficiency. J should be selected large enough to give an accurate estimate of p_0.

The SuS algorithm can be slightly modified for application to Bayesian updating, where $\Pr(Z_e)$ is not the main result. For Bayesian updating, one is interested in obtaining samples that fall into the domain Ω, i.e. samples conditional on Z_e. Therefore, we add a final step, which is the

generation of such samples. In this final step, K samples are generated, where K can be freely chosen. The resulting algorithm is presented in the following.

SuS-based algorithm for BUS

Define: J (number of samples in each intermediate step), K (number of final samples), p_0 (probability of intermediate subsets).

Sample from the original distribution:

1. Generate J samples $\mathbf{u}_1^{(j)}$, $j = 1, \dots, J$, from the $(n+1)$-variate independent standard normal distribution, φ_{n+1}.
2. Define the domain $\Omega_1 = \{H(\mathbf{u}) \leq b_1\}$, wherein b_1 is chosen as the p_0-percentile of the samples $H(\mathbf{u}_1^{(j)})$, $j = 1, \dots, J$.
3. $i = 1$

Sample from the conditional distributions:

4. Repeat while $b_i > 0$.
   a. $i = i + 1$
   b. Generate J conditional samples $\mathbf{u}_i^{(j)}$, $j = 1, \dots, J$, from the $(n+1)$-variate independent standard normal distribution conditional on Ω_{i-1}, $\varphi_{n+1}(\mathbf{u}|\Omega_{i-1})$, using a MCMC algorithm, e.g. (Au and Beck 2001) or (Papaioannou et al. 2014).
   c. Define the domain $\Omega_i = \{H(\mathbf{u}) \leq b_i\}$, wherein b_i is chosen as the p_0-percentile of the samples or 0, whichever is larger.

Sample from the posterior distribution:

5. Identify all samples from $\mathbf{u}_i$ that are in the domain $\Omega_U = \{H(\mathbf{u}) \leq 0\}$. Set Λ equal to the number of these samples
6. Generate K conditional samples $\mathbf{u}^{(k)}$, $k = 1, \dots, K$, from the $(n+1)$-variate independent standard normal distribution conditional on Ω_U, $\varphi_{n+1}(\mathbf{u}|\Omega_U)$. This uses a MCMC algorithm where the seeds are the Λ samples identified in 5.
7. Transform the samples $\mathbf{u}^{(k)}$ to the original space to obtain samples from the posterior distribution: $\mathbf{x}^{(k)} = \mathbf{T}\left(u_1^{(k)}, \dots, u_n^{(k)}\right)$, $k = 1, \dots, K$.

Estimate the acceptance probability:

8. $p_{acc} = \Pr(Z_e) \approx p_0^{M-1} \cdot \frac{\Lambda}{J}$, following Eq. (23), with $M = i$.

As pointed out earlier, the necessary computational effort is determined mainly by the number of likelihood function calls. In the above algorithm, the first and each intermediate step require J likelihood function calls; the final step (sampling from the posterior) requires an additional $K - \Lambda$ likelihood function calls. The total number of calls associated with the above algorithm is thus equal to $M \cdot J + K - \Lambda$, where M is the total number of steps. It is $M \approx \lfloor \log p_{acc} / \log p_0 \rfloor$, hence the computational effort is approximately proportional to $\log p_{acc}$.

Note that the number of likelihood function calls in the BUS approach is approximately equal to the number of limit state function calls in the traditional use of SuS (or other structural reliability methods), when evaluating the probability of an event with probability p_{acc}. For BUS, the fact that accuracy in computing p_{acc} is typically not crucial can motivate a reduction in the number of samples J in each step as compared to the original SuS, for which values of $J = 500$ and $p_0 = 0.1$ were proposed by (Au and Beck 2001). However, we do not investigate such further optimization in this paper.

Bayesian inference

Once K samples $\mathbf{x}^{(k)}$ from the posterior distribution are obtained, any property of this distribution can be estimated using a MCS approach. When the $\mathbf{x}^{(k)}$ are generated following the SuS algorithm outlined above, it must be taken into account that the samples $\mathbf{x}^{(k)}$ can be correlated (this is in analogy to MCMC). While this does not affect the estimates below, it does affect their accuracy.

The expected value of any function Y of $\mathbf{X}$ is estimated as

$$\mathrm{E}[Y(\mathbf{X})] \approx K^{-1} \sum_{k=1}^{K} Y(\mathbf{x}^{(k)}). \tag{24}$$

Accordingly, an estimate of the distribution of $Y(\mathbf{X})$ is obtained as

$$F_{Y(\mathbf{X})}(\mathrm{y}) = \Pr(Y(\mathbf{X}) \le y) = \mathrm{E}[I(Y(\mathbf{X}) \le y)] \approx K^{-1} \sum_{k=1}^{K} I\left[Y\left(\mathbf{x}^{(k)}\right) \le y\right]. \tag{25}$$

Alternatively, kernel density approximations may be used for the posterior of $\mathbf{X}$ or $Y(\mathbf{X})$ (Turlach 1993).

If the interest is in computing the posterior probability of a failure event F, defined through a limit state function $g(\mathbf{X})$ as $F = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \leq 0\}$, the corresponding function is $Y(\mathbf{X}) = I(g(\mathbf{X}) \leq 0)$. A MCS estimate from the posterior samples following Eq. (24) is inefficient when the posterior probability of failure is small. In such cases, which are common in structural reliability problems, it is beneficial to follow the procedure in (Straub 2011). Thereby, the intersection of F with the observation event Z_e is computed directly by means of structural reliability methods.

Illustrative applications

Three applications are presented. The first one demonstrates the applicability of BUS to problems that are not globally identifiable. The second demonstrates its applicability to problems for which a larger number of measurements with correlated errors are available. The third example demonstrates the applicability of the method to FE-based analysis, including a random field modeling of material (soil) properties.

Parameter identification in a two DoF system

This example is due to (Beck and Au 2002). The mechanical model, the prior distributions as well as the data are taken from the original reference, so that the results can be compared.

The example is a two degree of freedom (DoF) system, whose uncertain spring coefficients are to be determined based on measurements of the first two eigenfrequencies. The problem is not globally identifiable, i.e. there are multiple combinations of values of the model parameters that can well explain the measured eigenfrequencies. In (Beck and Au 2002), the problem was solved through Bayesian updating using a MCMC approach. Here, we employ the proposed BUS method implemented through the SuS-based algorithm.

A two-story frame-structure is modeled through a shear building model with two DoF (Figure 3). The story masses, which include the mass contributions from the columns, are taken as deterministic values with $m_1 = 16.531 \cdot 10^3$ kg and $m_2 = 16.131 \cdot 10^3$ kg. The inter-story stiffness values are modeled as $K_1 = X_1 k_n$ and $K_2 = X_2 k_n$, where $k_n = 29.7 \cdot 10^6$ N/m is the nominal value and X_1, X_2 are correction factors. Damping is not considered in the analysis. Observations of the first two eigenfrequencies f_1 and f_2 are used to update the prior distribution of $\mathbf{X} = [X_1; X_2]$ to its posterior distribution. Following (Beck and Au 2002), the likelihood function is

$$L(\mathbf{x}) \propto \exp\left[-\frac{J(\mathbf{x})}{2\sigma_\epsilon^2}\right], \tag{26}$$

where

$$J(\mathbf{x}) = \sum_{j=1}^{2} \lambda_j^2 \left[\frac{f_j^2(\mathbf{x})}{\tilde{f}_j^2} - 1\right]^2 \tag{27}$$

is a modal measure-of-fit function. $f_j^2(\mathbf{x})$ is the jth eigenfrequency predicted with the model with parameters $\mathbf{x}$ and $\tilde{f}_j^2$ is the corresponding measurement. $\lambda_1 = \lambda_2 = 1$ are the means and $\sigma_\epsilon = \frac{1}{16}$ is the standard deviation of the prediction error. The measured eigenfrequencies are $\widetilde{f_1} = 3.13\text{Hz}$ and $\widetilde{f_2} = 9.83\text{Hz}$.

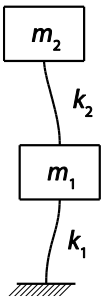

Figure 3. Two DoF shear building model.

The prior probability distributions of X_1, X_2 are uncorrelated lognormal distributions with modes 1.3 and 0.8 and standard deviations $\sigma_{X_1} = \sigma_{X_2} = 1$. The proportionality constant in the formulation of the likelihood function, Eq. (26), is selected as 1 and the constant c in the observation limit state function H is taken as 1. (The largest possible value of the likelihood function occurs for $J(\mathbf{x}) = 0$, in which case it is $L(\mathbf{x}) = 1$. Therefore, with $c = 1$ it is $cL(\mathbf{x}) \leq 1$ for all $\mathbf{x}$.)

The samples obtained with the SuS-based rejection sampling algorithm are shown in Figure 4, including the samples from the intermediate steps. The parameters of the algorithm were selected as $J = 500$, $K = 500$ and $p_0 = 0.1$, without attempting any optimization. The number of subset simulation steps is $M = 3$ and the number of additional samples in the last step is $K - \Lambda = 349$. It follows that the total number of samples, and hence the total number of likelihood function evaluations, is 1849.

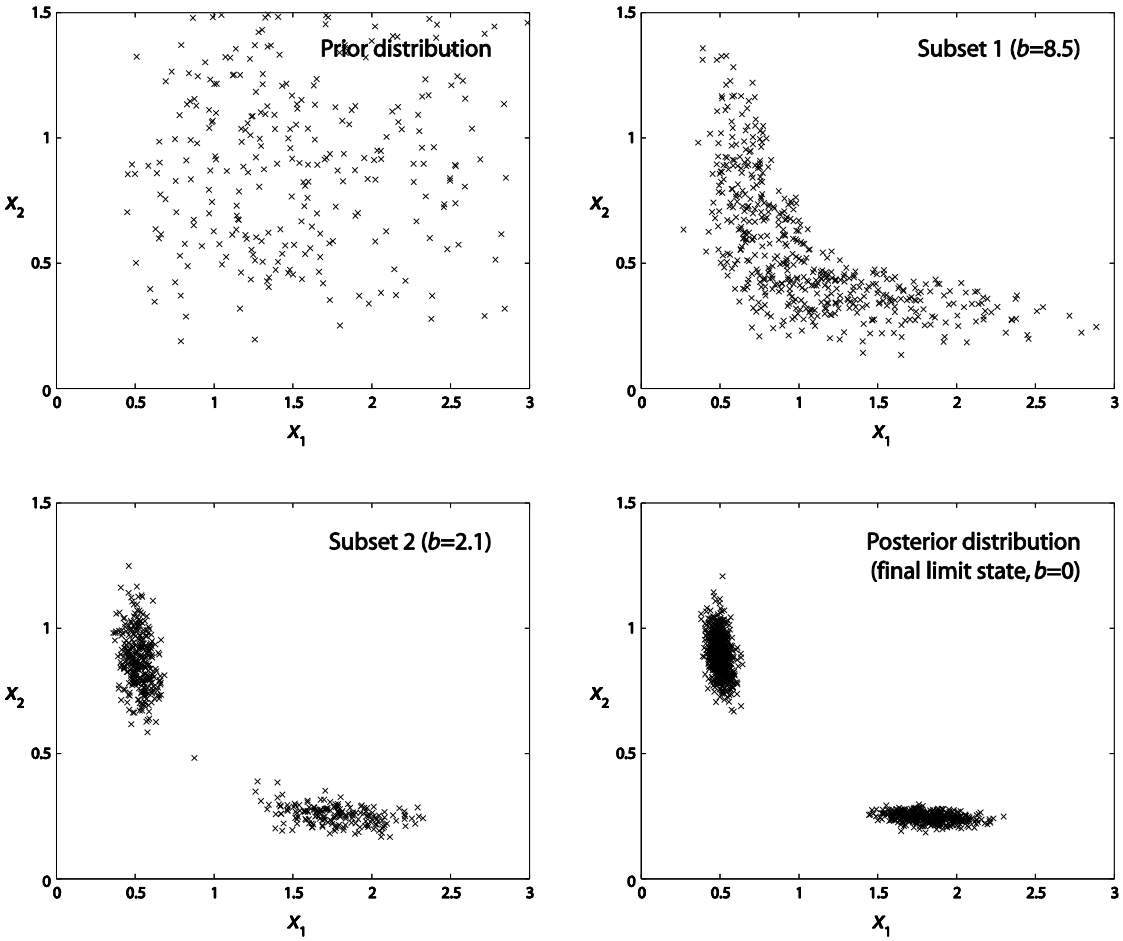

Figure 4. Bayesian updating of the correction factors $\boldsymbol{X} = [X_1; X_2]$. Samples from the prior distribution, the intermediate subsets and the final posterior distribution, shown in the outcome space of $\boldsymbol{X}$. In each step, 500 samples were generated, but some samples are coinciding.

In Figure 5, the resulting posterior marginal CDF of X_1 is shown for several randomly selected runs of the algorithm. The bi-modal nature of the posterior distribution, which is already evident from Figure 4, can be clearly observed also in the marginal CDF. The comparison with the exact result reveals that the shape of the distribution around the two modes is well captured by the samples. The probability of the two modes is estimated with less accuracy, as evident from the scatter in the value of the CDF between 0.5 and 1.5. The same observation is made in (Beck and Au 2002). This scatter could be reduced by increasing the number of samples; alternatively, additional importance sampling evaluations may help to reduce this scatter. For most practical applications the accuracy of the presented results would be sufficient. The more important result is that the proposed method can successfully identify both modes of the posterior distribution.

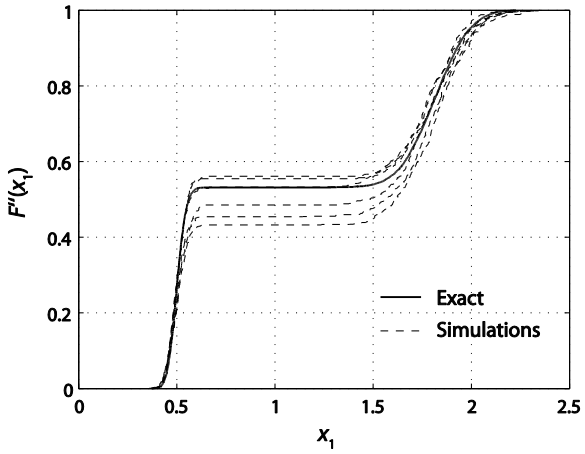

Figure 5. Posterior CDF of X_1 obtained from 6 repeated runs of the proposed algorithm, together with the exact result.

The results shown in Figure 4 and Figure 5 are consistent with the results given in (Beck and Au 2002), where further details and discussion of the example can be found.

Bayesian updating of the flexibility of a cantilever beam

This example demonstrates the application of the proposed method to learning a random field with measurements whose errors are correlated. The example has an analytical solution, which allows a validation of the method.

We update the spatially variable flexibility $F(x)$ of a cantilever beam based on measurements of the beam deflections. The beam has length $L = 5\text{m}$ and is subjected to a deterministic point load $V = 20\text{kN}$ at the free end (Figure 6). The flexibility $F(x)$ is defined as the reciprocal of the bending rigidity of the beam, so that $F(x) = 1/E(x)I$, where $E(x)$ is the Young's modulus at location x and I is the moment of inertia. We assume that the prior distribution of $F(x)$ is described by a homogeneous Gaussian random field with mean $\mu_F = 10^{-7}\text{N}^{-1}\text{m}^{-2}$ and exponential auto-covariance function $\Gamma_{FF}(x_1, x_2) = \sigma_F^2 \exp\left(-\frac{|x_1 - x_2|}{\lambda}\right)$, where $\sigma_F = 3.5 \times 10^{-8}\text{N}^{-1}\text{m}^{-2}$ is the standard deviation and $\lambda = 2\text{m}$ the correlation length.

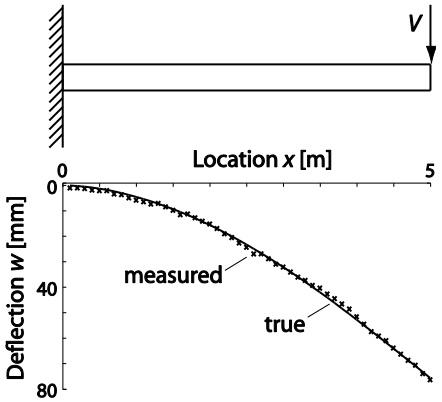

Figure 6. Beam subjected to point load V; measured deflection.

The second derivative of the vertical deflection $W(x)$ is a function of $F(x)$ and the bending moment $M(x)$:

$$\frac{d^2W(x)}{dx^2} = -M(\mathrm{x})F(x) = V(L-x)F(x)\,. \tag{28}$$

Taking the double integral of the above, it is:

$$W(x) = \int_0^{\mathrm{x}} \int_0^{s} V(L-t)F(t)\,\mathrm{d}t\,\mathrm{d}s\,. \tag{29}$$

Using the prior information on $F(x)$, one can evaluate the prior mean and auto-covariance function of $W(x)$. Moreover, because $F(x)$ is Gaussian and $W(x)$ is a linear function of $F(x)$, the prior distribution of $W(x)$ will also be Gaussian. The mean of $W(x)$ reads:

$$\mu_w(x) = \int_0^{\mathrm{x}} \int_0^{s} V(L-t)\mu_F\,\mathrm{d}t\,\mathrm{d}s = \frac{\mu_F}{6}x^2(3L-x) \tag{30}$$

The auto-covariance function of $W(x)$ is:

$$\begin{aligned}\Gamma_{ww}(x_1,x_2) &= \int_0^{\mathrm{x}_2}\int_0^{x_1}\int_0^{s_2}\int_0^{s_1} V^2(L-t_1)(L-t_2)\Gamma_{FF}(t_1,t_2)\,\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}s_1\,\mathrm{d}s_2 \\ &= \int_0^{\mathrm{x}_2}\int_0^{x_1}\int_0^{s_2}\int_0^{s_1} V^2(L-t_1)(L-t_2)\sigma_F^2 \exp\left(-\frac{|t_1-t_2|}{\lambda}\right)\mathrm{d}t_1\,\mathrm{d}t_2\,\mathrm{d}s_1\,\mathrm{d}s_2\,.\end{aligned} \tag{31}$$

The above integral can be solved analytically, however the resulting expression is lengthy and hence omitted here.

Measurements $\mathbf{w}_m$ of the deflection are made at 50 points (0.1m, 0.2m, …, 5m) along the beam using optic measurements. The measurements are subjected to additive errors, which are described by a joint normal PDF $f_{\boldsymbol{\epsilon}}$ with zero mean and covariance matrix $\mathbf{C}_{\boldsymbol{\epsilon\epsilon}}$, whose elements are determined from the auto-covariance function $\Gamma_{\epsilon\epsilon}(x_1, x_2) = \sigma_\epsilon^2 \exp\left(-\frac{|x_1 - x_2|}{\lambda_\epsilon}\right)$ with $\lambda_\epsilon = 1\text{m}$. The standard deviation of the measurement error is $\sigma_\epsilon = 1\text{mm}$. The simulated measurements are shown in Figure 6, together with the true (but in real applications unknown) deflection of the beam.

The likelihood function describing these measurements is

$$L(\mathbf{f}) = f_{\boldsymbol{\epsilon}}[\mathbf{w}(\mathbf{f}) - \mathbf{w}_m] \tag{32}$$

where $\mathbf{f} = [f(0.1\text{m}); f(0.2\text{m}); \ldots; f(5\text{m})]$ is a vector describing the flexibility at the 50 discretization points and $\mathbf{w}(\mathbf{f})$ is the vector of deflections at these 50 locations computed from $\mathbf{f}$ following Eq. (29).

Bayesian updating of the flexibility at the locations $\mathbf{x} = [0.1\text{m}; 0.2\text{m}; \ldots; 5\text{m}]$ is performed with BUS using the SuS-based algorithm with $J = 10^3$, $K = 5 \times 10^3$ and $p_0 = 0.1$. The constant c is selected as the inverse of the likelihood function value at the MLE following Annex A. The resulting number of subsets is $M = 4$.

Estimates of the flexibility based on the deflection measurements are presented in Figure 7. The results show the advantage of a Bayesian analysis over MLE for this problem. Due to the fact that 50 (correlated) measurements are available for estimating 50 parameters, the problem is ill-posed and MLE leads to overfitting. The Bayesian analysis with informative prior distribution regularizes the problem and provides a good approximation to the true values.

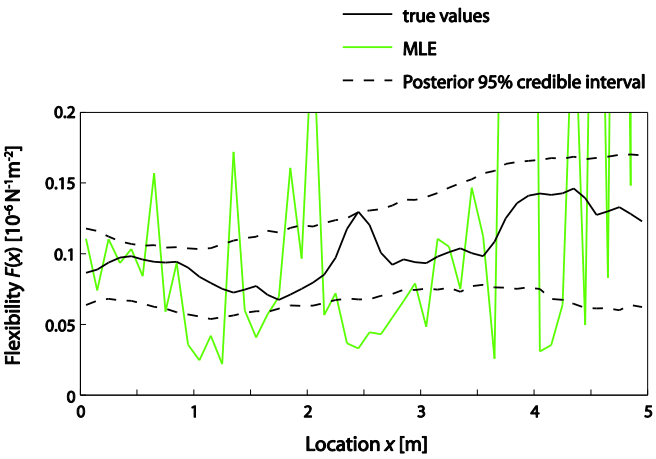

Figure 7. Posterior credible interval of the flexibility computed with BUS, together with the true value and the maximum likelihood estimate (MLE).

Figure 8 presents the deflection of the beam, comparing the true value, the measured value, the MLE and the Bayesian estimate. Figure 8a illustrates that the measured values are very similar in absolute terms. To enable an appraisal of the differences in the estimates, Figure 8b presents the difference of the deflections from the ones computed with the prior mean of $F(x)$. The Bayesian analysis provides a good estimate of the actual deflections, whereas the MLE leads to an overfitting to the measurements in agreement with the results of Figure 7.

Since the prior distribution of $W(x)$ is Gaussian and the measurement errors are additive and also jointly Gaussian, an analytical solution of the posterior joint PDF of $W(x)$ can be obtained for validating the results obtained with BUS. The results in Figure 8 show that the posterior credible intervals obtained with BUS coincide almost perfectly with those calculated analytically. Further proof of the accuracy of BUS is available from Figure 9, which shows a comparison of the analytically calculated posterior CDF of the deflection at the free end with the BUS solution.

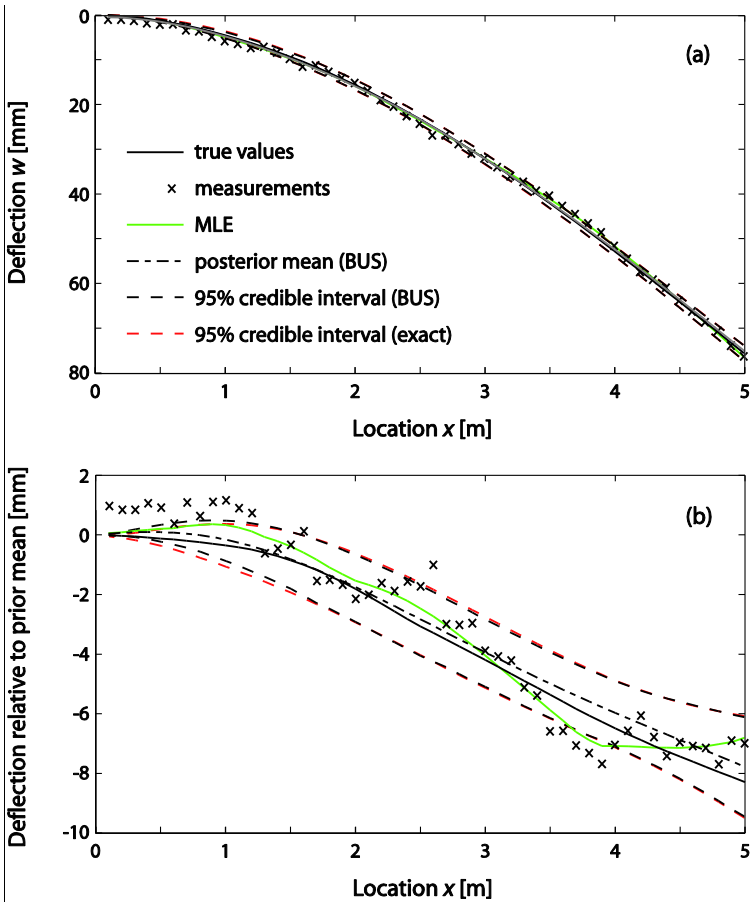

Figure 8. (a) Posterior credible interval of the deflection, together with the true value, the measurements and the maximum likelihood estimate (MLE). (b) As in (a), but showing the difference of the deflection to its prior mean.

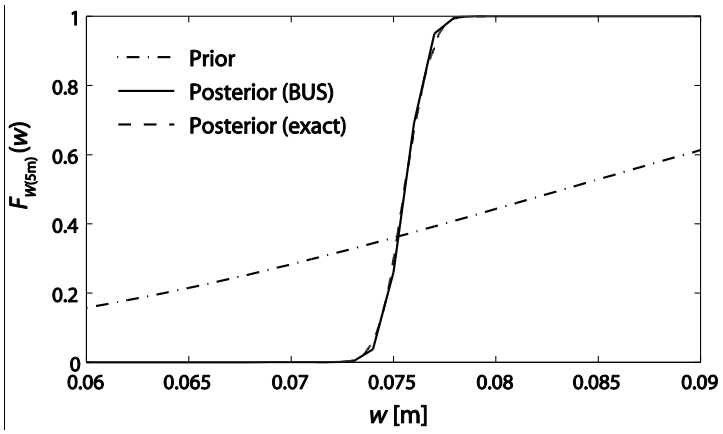

Figure 9. Marginal CDF of the deflection at the end of the cantilever $W(5m)$. Prior CDF, together with the posterior CDF computed using the proposed method and the analytical solution.

Finite element model updating of a geotechnical construction

This example is based on previous reliability analysis and updating presented in (Papaioannou and Straub 2012). We update the material properties of the soil surrounding a geotechnical site based on a deformation measurement performed in situ.

The site consists of a 5.0m deep trench with cantilever sheet piles in a homogeneous soil layer of dense cohesionless sand with uncertain spatially varying mechanical properties (see Figure 10). The soil is modelled in 2D with plane-strain finite elements. For simplicity, neither groundwater nor external loading is considered. Additionally, we take advantage of the symmetry of the trench and model just half of the soil profile, although this implies an approximation when randomness in the soil material is taken into account. The material model used is an elasto-plastic model with a prismatic yield surface according to the Mohr-Coulomb criterion and a non-associated plastic flow. The sheet pile is modelled using beam elements and the interaction between the retaining structure and the surrounding soil is modelled using nonlinear interface elements. The corresponding FE model is implemented in the SOFiSTiK program (SOFiSTiK 2012). Further details on the mechanical model and the simulation of the excavation process can be found in (Papaioannou and Straub 2012).

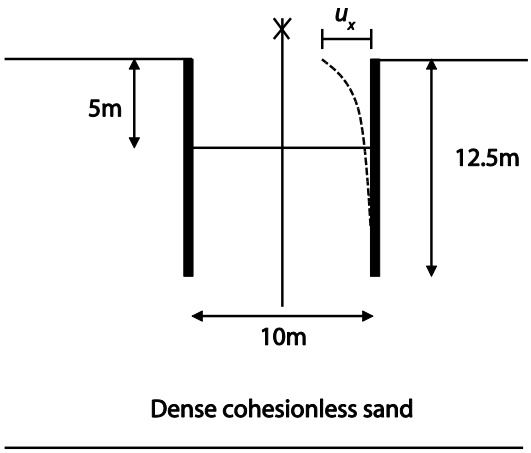

Figure 10. Sheet pile wall in sand.

Homogeneous non-Gaussian random fields describe the prior distributions of the uncertain material properties: Young's modulus E, friction angle ϕ and specific weight γ. The joint distribution at each pair of locations is modelled by the Nataf distribution (Kiureghian and Liu 1986) with marginal distributions according to Table 1. The auto-correlation coefficient function is given by a separable exponential model $\rho_{XZ}(\tau_x, \tau_z) = \exp(-\frac{\tau_x}{\lambda_x} - \frac{\tau_z}{\lambda_z})$, where τ_x, τ_z are the absolute distances in the x (horizontal) and z (vertical) directions. The correlation lengths are $\lambda_x = 20\text{m}$ and $\lambda_z = 5\text{m}$ for all uncertain soil material properties. Cross-correlation between the different material properties is not included. The random fields are discretized by the midpoint method (Der Kiureghian and Ke 1988) using a stochastic mesh, consisting of 144 deterministic FE patches. The stochastic discretization resulted in a total of $3 \times 144 = 432$ basic random variables gathered in a vector **X**. In Figure 11, the stochastic and deterministic FE meshes are shown. Figure 11c shows the deformed configuration at the final excavation stage computed with the mean values of the random fields.

Table 1. Prior marginal distributions of the material properties of the soil.

Parameter	Distribution	Mean	COV
Specific weight γ [kN/m^3]	Normal	19.0	5%
Young's modulus E [MPa]	Lognormal	125.0	25%
Poisson's ratio ν	-	0.35	-
Friction angle φ [°]	Beta(0.0,45.0)	35.0	10%
Cohesion c [MPa]	-	0.0	-
Dilatancy angle ψ [°]	-	5.0	-

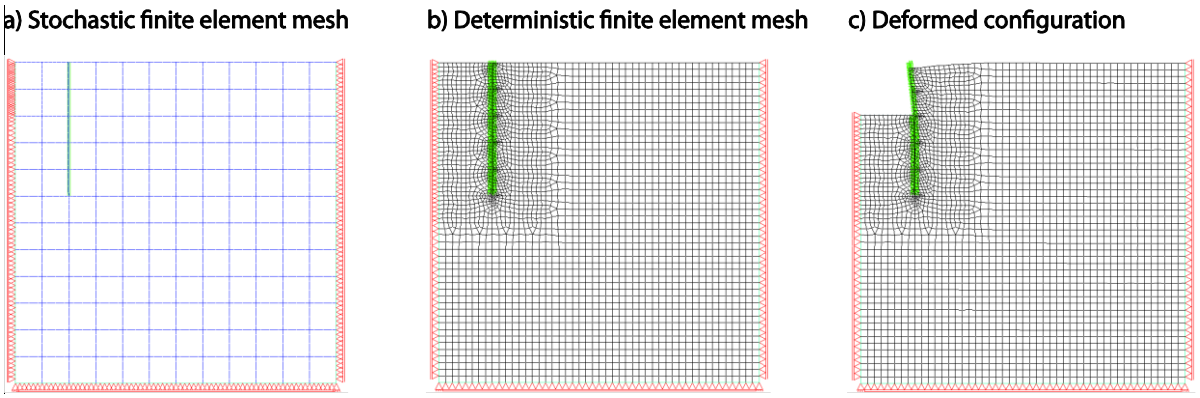

Figure 11. (a) Stochastic and (b) deterministic finite element mesh of the geotechnical site, shown for the situation prior to the excavation; (c) deformed configuration at full excavation for the mean values of the material properties.

We assume that a measurement of the horizontal displacement at the top of the trench $u_{x,m} = 60\text{mm}$ is made at full excavation. The measurement is subjected to an additive error ϵ, which is described by a normal PDF f_ϵ with zero mean and standard deviation $\sigma_\epsilon = 5\text{mm}$. The likelihood function describing the measurement is

$$L(\mathbf{x}) = f_\epsilon\left[u_{x,m} - u_x(\mathbf{x})\right] \tag{33}$$

where $\mathbf{x}$ describes the material properties at the midpoints of the stochastic elements and $u_x(\mathbf{x})$ is the displacement evaluated by the FE program. Bayesian updating of the vector $\mathbf{X}$ is performed with BUS using the SuS-based algorithm with $J = 10^3$, $K = 1.5 \times 10^3$ and $p_0 = 0.1$. The constant c is selected as $c = \sigma_\epsilon$, which satisfies the condition $cL(\mathbf{x}) \leq 1$.

The prior mean of $u_x(\mathbf{X})$ is 50.2mm, which indicates that the prior model underestimates the measured tip displacement. Figure 12 and Figure 13 show the posterior mean of the Young's modulus E and friction angle ϕ, respectively. The posterior means of the elements in the vicinity of the trench are smaller than the prior mean, which reflects the effect of the measured displacement. The effect is local, since the values of the stiffness and strength of the soil farther away from the trench have limited influence on the deformation at the location of the measurement. Moreover, the results show the influence of the auto-correlation of the prior distribution. The change of the posterior means is steeper in the vertical than in the horizontal direction, which is due to the fact that the prior correlation length in the horizontal direction is larger than in the vertical direction. The low values of the Young's modulus observed in the bottom right of Figure 12 cannot be explained by the measurement and hence are attributed to sampling error.

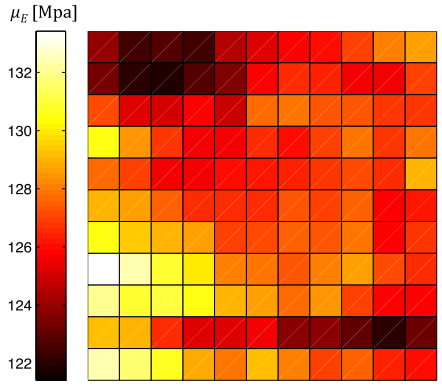

Figure 12. Posterior mean of the Young's modulus E of the soil.

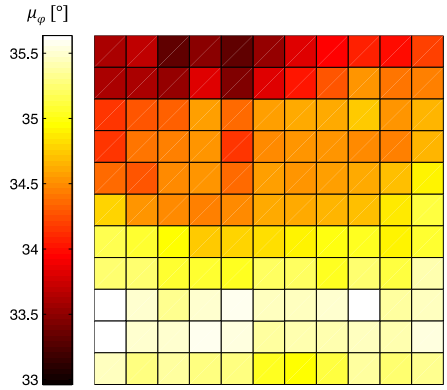

Figure 13. Posterior mean of the friction angle ϕ of the soil.

Discussion

Bayesian updating using structural reliability methods (BUS) has been developed for learning and calibrating engineering and computer models. By interpreting Bayesian updating as a structural reliability problem, the existing suite of methods for solving such problems is available, including importance sampling, subset simulation and first- and second-order reliability methods (FORM/SORM). For analysts with experience in these methods, it is straightforward to implement the BUS approach.

The presented implementation of BUS through the subset simulation (SuS) algorithm is ready-to-use for many practically relevant applications. The presented numerical applications as well as further numerical investigations not reported in this paper showed that the SuS-based algorithm works robustly and efficiently for a broad range of problems. It also has the advantage of being applicable in high dimensions – at least theoretically its performance does not decrease with increasing number of random variables. It is pointed out that the algorithm uses MCMC

to generate the conditional samples required by SuS. Since the initial samples of these Markov chains are already from the target distribution, it is ensured that all samples follow this target distribution and convergence of the chain is not an issue. However, the MCMC procedure does introduce a correlation among the resulting samples, which must be taken into account when evaluating the posterior statistics.

BUS has originally been developed for problems with an informative prior, i.e. when relevant information on the uncertain variables is available a-priori. This is a common situation when learning mechanical models, where significant prior knowledge on uncertain parameters is often available. But BUS is a viable alternative to existing methods also for problems with weakly or non-informative prior distributions (e.g. in parameter identification), as demonstrated by application 1 in this paper.

In this work, no attempt was made to optimize the structural reliability methods for their use in BUS. We are currently investigating further optimizations of SuS targeted towards its implementation in BUS, including an adaptive choice of the constant c. In addition, we believe that there is a significant potential in identifying alternative structural reliability methods that can be more efficient than SuS in combination with BUS for specific types of applications. As an example, the use of importance sampling methods around the design point (which in BUS corresponds to the mode of the posterior distribution) appears promising, but has not been considered here.

In the applications presented in this paper, the number of model evaluations required for Bayesian updating is on the order of $10^3 - 10^4$, which is similar to existing state-of-the-art methods such as transitional MCMC proposed in (Ching and Chen 2007) or the hybrid MCMC approach of (Cheung and Beck 2009). While the presented implementation of BUS certainly offers room for a further enhancement of the computational efficiency, we do not believe that this number can be significantly reduced with any purely sampling-based approach. For expensive computational models, the BUS approach can be combined with surrogate models (response surfaces) to further reduce the number of model evaluations. Experience with adaptive response surfaces available in the structural reliability community may be used for this purpose.

In many instances, one is interested in updating the reliability of a mechanical system, or more generally in updating the probability of a rare event F with observations. Unlike other methods presented in the literature, BUS does not require that one first updates the joint probability distribution of the relevant variables $\mathbf{X}$ and then computes the updated probability of F based

on an approximation of the posterior PDF of **X**. With BUS, the observation is represented by the event Z_e and the posterior probability of F is equal to $\Pr(F|Z_e) = \Pr(F \cap Z_e) \,/\, \Pr(Z_e)$, see (Straub 2011). Therefore, Bayesian updating reduces to solving two structural reliability problems.

Conclusion

A novel approach to Bayesian updating of mechanical and general engineering models was proposed, termed BUS. It is based on interpreting the Bayesian updating problem as a structural reliability problem. This enables one to use the whole portfolio of existing structural reliability methods for performing Bayesian updating. In this paper, subset simulation was used to obtain samples of the posterior distribution. Three application examples were included to demonstrate the versatility and efficiency of BUS.

Annex A: How to select the constant c defining the observation domain?

BUS requires the selection of the constant c in Eq. (10). On the one hand, it is required that $cL(\mathbf{x}) \leq 1$ for any $\mathbf{x}$. On the other hand, the acceptance rate is directly proportional to c; i.e. the larger c, the larger $\Pr([\mathbf{X}, P] \in \Omega)$, which is beneficial for the MCS-based rejection sampling as well as for BUS in combination with subset simulation. It follows that an optimal choice of c is

$$c = \frac{1}{\sup L(\mathbf{x})}. \tag{34}$$

In some instances, $\sup L(\mathbf{x})$ is directly available. In case of pure inequality observations, it is $\sup L(\mathbf{x}) = 1$ and it follows that $c = 1$. In case of a single measurement with error ϵ, $\sup L(\mathbf{x})$ is equal to maximum of the PDF of ϵ, which is readily available. However, it is not always straightforward to evaluate Eq. (34) when there are several observations with corresponding likelihoods $L_i(\mathbf{x}), i = 1, \dots, m$. $\sup L(\mathbf{x})$ occurs at the location of the maximum likelihood, so if it is possible to determine the maximum likelihood estimator $\mathbf{x}_{MLE}$ using an optimization algorithm, it is $\sup L(\mathbf{x}) = L(\mathbf{x}_{MLE})$. In case $\mathbf{x}_{MLE}$ is not available, an alternative suboptimal choice is

$$c = \frac{1}{\prod_{i=1}^{m} [\sup L_i(\mathbf{x})]}. \tag{35}$$

For large m, i.e. when many individual observations are available, this choice is inefficient. In these situations, one can get an estimate of the statistical properties of $\sup L(\mathbf{x})$ to obtain a more efficient value of c, as demonstrated through an example in the following.

Consider a case where all measurements y_i have random errors ϵ_i that are independent identically distributed (iid) following a normal distribution with mean zero and standard deviation σ_ϵ. The likelihood function is

$$\begin{aligned} L(\mathbf{x}) &= \prod_{i=1}^{m} L_i(\mathbf{x}) \\ &= \prod_{i=1}^{m} \exp\left\{-\frac{1}{2}\frac{[h_i(\mathbf{x}) - y_i]^2}{\sigma_\epsilon^2}\right\} \\ &= \exp\left\{-\frac{1}{2}\frac{1}{\sigma_\epsilon^2}\sum_{i=1}^{m}[h_i(\mathbf{x}) - y_i]^2\right\}. \end{aligned} \tag{36}$$

Let us assume that the model is perfect if the true parameters $\mathbf{x}_0$ are used, i.e. $h_i(\mathbf{x}_0)$ are the true values of the measured quantities. In this case, the actual measurements y are $y_i = h_i(\mathbf{x}_0) + \epsilon_i$, i.e. they are normal iid random variables with means $h_i(\mathbf{x}_0)$ and standard deviation σ_ϵ. The maximum of the log-likelihood is then approximately

$$\begin{aligned} \max \ln L(\mathbf{x}) &\approx \ln L(\mathbf{x}_0) \\ &= -\frac{1}{2}\frac{1}{\sigma_\epsilon^2}\sum_{i=1}^{m}[h_i(\mathbf{x}_0) - y_i]^2 \\ &= -\frac{1}{2}\frac{1}{\sigma_\epsilon^2}\sum_{i=1}^{m}\epsilon_i^2 \\ &= -\frac{1}{2}\sum_{i=1}^{m}U_i^2, \end{aligned} \tag{37}$$

where U_i are independent standard normal random variables. The approximation becomes better with increasing m. It follows from Eq. (37) that $\{-2 \max \ln L(\mathbf{x})\}$ is approximately chi-squared distributed with m degrees of freedom. Therefore, the distribution of the maximum likelihood is independent of the observation error ϵ_i and depends only on the number m of independent measurements. One can thus select a value of c, for which it holds $cL(\mathbf{x}) \leq 1$ with given probability p, as

$$c = \frac{1}{\exp\left[-\frac{1}{2}F_{X^2(m)}^{-1}(p)\right]}. \tag{38}$$

where $F_{X^2(m)}^{-1}$ is the inverse CDF of the chi-squared distribution with m degrees of freedom. E.g. if the number of independent measurements is $m = 50$, a value $c = 3.5 \times 10^7$ will ensure that $cL(\mathbf{x}) \leq 1$ for all $\mathbf{x}$ with probability $p = 0.05$.

In reality, the model will not be perfect, which will lead to additional deviations of the model prediction from the observations. Therefore, the maximal log-likelihood will be lower than predicted by Eq. (37), and the probability that c is too large will be smaller than p.

Annex B: Probability of accepting samples in the simple rejection sampling algorithm

To illustrate the dependence of the average acceptance rate of samples p_{acc} on the number of observations, we consider m measurements of the quantity X, all with iid observation errors. This is in analogy to the illustration of the simple rejection sampling algorithm provided in the paper.

The likelihood function with m observations y_i, $i = 1, \dots, m$, with iid normal additive error with zero mean is:

$$L(x) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{m}\frac{(x-y_i)^2}{\sigma_\epsilon^2}\right] \tag{39}$$

The measurement values can be written as $y_i = x_0 + \epsilon_i = x - \Delta x + \epsilon_i$, where x_0 is the true (but unknown) value of X, ϵ_i is the observation error and $\Delta x = x_0 - x$ is the distance between x and the unknown true value of X. Inserting this relationship in (39), setting the proportionality constant to 1 and taking logarithms, one obtains

$$\begin{aligned}\ln L(x) &= -\frac{1}{2}\sum_{i=1}^{m}\frac{(\Delta x-\epsilon_i)^2}{\sigma_\epsilon^2} \\ &= -\frac{1}{2}\frac{1}{\sigma_\epsilon^2}\sum_{i=1}^{m}\Delta x^2 - 2\epsilon_i\Delta x + \epsilon_i^2.\end{aligned} \tag{40}$$

The expected value with respect to the observation errors ϵ_i is

$$\mathrm{E}[\ln L(x)] = -\frac{1}{2}\frac{1}{\sigma_\epsilon^2}(m\Delta x^2 + m\mathrm{E}[\epsilon_i^2]) = -\frac{m}{2}\left(\frac{\Delta x^2}{\sigma_\epsilon^2}+1\right). \tag{41}$$

The first-order approximation of the expected value of the likelihood is

$$\mathrm{E}[L(\mathrm{x})] \approx \exp\left[-\frac{m}{2}\left(\frac{\Delta x^2}{\sigma_\epsilon^2}+1\right)\right]. \tag{42}$$

The width of this likelihood function is proportional to $1/\sqrt{m}$. It follows that the expected number of samples falling into this domain, and consequently p_{acc}, is also proportional to $1/\sqrt{m}$.

References

Au, S.-K. and J. L. Beck (2001). "Estimation of small failure probabilities in high dimensions by subset simulation." Probabilistic Engineering Mechanics **16**(4): 263-277.

Beck, J. and L. Katafygiotis (1998). "Updating Models and Their Uncertainties. I: Bayesian Statistical Framework." Journal of Engineering Mechanics **124**(4): 455-461.

Beck, J. L. and S. K. Au (2002). "Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation." Journal of Engineering Mechanics-Asce **128**(4): 380-391.

Benjamin, J. and C. A. Cornell (1970). Probability, Statistics, and Decision for Civil Engineers, McGraw-Hill, New York.

Beven, K. (2008). Environmental modelling: an uncertain future?, Taylor & Francis.

Bolstad, W. M. (2011). Understanding computational Bayesian statistics, Wiley.

Bucher, C. G. (1988). "Adaptive sampling — an iterative fast Monte Carlo procedure." Structural Safety **5**(2): 119-126.

Cheung, S. H. and J. L. Beck (2009). "Bayesian Model Updating Using Hybrid Monte Carlo Simulation with Application to Structural Dynamic Models with Many Uncertain Parameters." Journal of Engineering Mechanics-Asce **135**(4): 243-255.

Ching, J. and Y. Chen (2007). "Transitional Markov Chain Monte Carlo Method for Bayesian Model Updating, Model Class Selection, and Model Averaging." Journal of Engineering Mechanics **133**(7): 816-832.

Chopin, N. (2002). "A sequential particle filter method for static models." Biometrika **89**(3): 539-552.

Der Kiureghian, A. and J.-B. Ke (1988). "The stochastic finite element method in structural reliability." Probabilistic Engineering Mechanics **3**(2): 83-91.

Der Kiureghian, A. and P.-L. Liu (1986). "Structural reliability under incomplete probability information." Journal of Engineering Mechanics **112**(1): 85-104.

Ditlevsen, O. and H. O. Madsen (1996). Structural reliability methods. Chichester [u.a.], Wiley.

Doucet, A., N. De Freitas and N. Gordon (2001). Sequential Monte Carlo methods in practice, Springer New York.

Fisher, R. A. (1922). "On the Mathematical Foundations of Theoretical Statistics." Philosophical Transactions of the Royal Society of London. Series A **222**(594-604): 309-368.

Gelman, A. (2004). Bayesian data analysis. Boca Raton, Fla., Chapman & Hall/CRC.

Gilks, W. R., S. Richardson and D. J. Spiegelhalter (1998). Markov chain Monte Carlo in practice. Boca Raton, Fla., Chapman & Hall.

Gilks, W. R., A. Thomas and D. J. Spiegelhalter (1994). "A language and program for complex Bayesian modelling." The Statistician: 169-177.

Gilks, W. R. and P. Wild (1992). "Adaptive rejection sampling for Gibbs sampling." Applied Statistics **41**(2): 337-348.

Hohenbichler, M. and R. Rackwitz (1981). "Non-Normal Dependent Vectors in Structural Safety." Journal of the Engineering Mechanics Division-Asce **107**(6): 1227-1238.

Kavetski, D., G. Kuczera and S. W. Franks (2006). "Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory." Water Resources Research **42**(3): n/a-n/a.

Kennedy, M. C. and A. O'Hagan (2001). "Bayesian calibration of computer models." Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63**(3): 425-464.

Kiureghian, A. D. and P. L. Liu (1986). "Structural Reliability under Incomplete Probability Information." Journal of Engineering Mechanics-Asce **112**(1): 85-104.

Kurtz, N. and J. Song (2013). "Cross-entropy-based adaptive importance sampling using Gaussian mixture." Structural Safety **42**(0): 35-44.

Madsen, H. O., S. Krenk and N. C. Lind (1985). Methods of structural safety. Englewood Cliffs, NJ, Prentice-Hall.

Mahadevan, S., R. Zhang and N. Smith (2001). "Bayesian networks for system reliability reassessment." Structural Safety **23**(3): 231-251.

Natke, H. G. (1988). "Updating computational models in the frequency domain based on measured data: a survey." Probabilistic Engineering Mechanics **3**(1): 28-35.

Papadimitriou, C., J. L. Beck and L. S. Katafygiotis (2001). "Updating robust reliability using structural test data." Probabilistic Engineering Mechanics **16**(2): 103-113.

Papaioannou, I., W. Betz, K. Zwirglmaier and D. Straub (2014). "MCMC algorithms for subset simulation." Manuscript, Engineering Risk Analysis Group, Technische Universität München, Germany.

Papaioannou, I. and D. Straub (2012). "Reliability updating in geotechnical engineering including spatial variability of soil." Computers and Geotechnics **42**: 44-51.

Plummer, M., N. Best, K. Cowles and K. Vines (2006). "CODA: Convergence diagnosis and output analysis for MCMC." R news **6**(1): 7-11.

Rackwitz, R. (2001). "Reliability analysis - a review and some perspectives." Structural Safety **23**(4): 365-395.

Smith, A. F. M. and A. E. Gelfand (1992). "Bayesian Statistics without Tears: A Sampling-Resampling Perspective." The American Statistician **46**(2): 84-88.

SOFiSTiK (2012). SOFiSTiK analysis programs. Version 2012. Oberschleissheim, SOFiSTiK AG.

Straub, D. (2009). "Stochastic Modeling of Deterioration Processes through Dynamic Bayesian Networks." Journal of Engineering Mechanics-Asce **135**(10): 1089-1099.

Straub, D. (2011). "Reliability updating with equality information." Probabilistic Engineering Mechanics **26**(2): 254-258.

Straub, D. and A. Der Kiureghian (2010a). "Bayesian Network Enhanced with Structural Reliability Methods: Application." Journal of Engineering Mechanics - ASCE **136**(10): 1259-1270.

Straub, D. and A. Der Kiureghian (2010b). "Bayesian Network Enhanced with Structural Reliability Methods: Methodology." Journal of Engineering Mechanics-ASCE **136**(10): 1248-1258.

Straub, D. and M. H. Faber (2005). "Risk based inspection planning for structural systems." Structural Safety **27**(4): 335-355.

Sundar, V. S. and C. S. Manohar (2013). "Updating reliability models of statically loaded instrumented structures." Structural Safety **40**(0): 21-30.

Tang, W. (1973). "Probabilistic updating of flaw information." Journal of Testing and Evaluation **1**: 459-467.

Turlach, B. A. (1993). Bandwidth selection in kernel density estimation: A review. Discussion Paper 9307, Institut für Statistik und Ökonometrie, Humboldt-Universität zu Berlin.

Yang, J. and W. Trapp (1974). "Reliability Analysis of Aircraft Structures under Random Loading and Periodic Inspection." Aiaa Journal **12**: 1623-1630.

Zuev, K. M., J. L. Beck, S.-K. Au and L. S. Katafygiotis (2012). "Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions." Computers & Structures **92**: 283-296.